DiffuserCam: lensless single-exposure 3D imaging using Custom Hardware

Ray Duran  
Department of Electrical Engineering and Computer Science  
Univertisy of North Dakotaline 4: City, Countryray.duran@und.edu

*Abstract*— **In the paper by Antipa et al. [1], a lens less diffuser cam for 3-D imaging is introduced. Using a unique pseudorandom pattern of caustics on the sensor, simple calibration, and computational processing we solve for 3-D voxels with a single exposure. The use of compressive sensing and a numerical method allows us to solve the matrix inversion problem without dealing with the large original matrix. The simplified convolution forward model, validated experimentally, has a FoV of +/- 42 degrees in the x -axis and +/- 30.5 degrees in y-axis and is object dependent for resolution as reported in [2] (Supplementary Antipa). We are constructing a custom enclosure for our CMOS sensor that includes a diffuser and aperture to duplicate the results in [1]. We also followed the instructions in a tutorial reference by [1], where we used commodity parts and homemade optics where we were not able to duplicate experimental results.**

# Background

The main motivation for a DiffuserCam is to provide a light field representation of 4-D information into a 2-D sensor. While capturing spatial and angle information we want the system to be inexpensive, flexible, and compact. Providing such a system as described above falls in the realm of light-field imaging which lends itself to many applications: 3D neural activity[3]. (ref: Pegard et. Al), compressive radar imaging[4] (ref: Richard Baraniuk), synthetic aperture[5] (ref from Cai paper ref 9), and visual odometry[6](ref in Cai ref 11)

Why diffusers for lensless systems? At first glance a diffuser is a highly diffractive medium that would not seem to lend itself to light field imaging. A coded aperture light field is an alternative to a diffuser that can also achieve higher resolution than conventional cameras [7] (ref: JIn paper ref 2). However, an amplitude mask system by its’ very nature limits the amount of light that the sensor can record. A micro lens array can gather more information that just x-y coordinates but suffers from scaling.

One of the key theoretical pillars in being able to use diffusers in a lensless system is that the surface of the diffuser can be modeled as a smooth Gaussian smooth surface. So the diffractive effect causing speckles (caustic patterns) from interference can be used for reconstruction. This is because each area of the diffuser is unique creating a type of signature that can encode plenoptic information about our illuminated object. These properties, of uniqueness, in theory allow us a better scaling that can be utilized by a micro lens array system. Another important point is that diffusers which are essentially phase shifter elements can concentrate the light from a 4-D light field into a 2-D sensor better than an amplitude mask using Fourier optics.

For the reasons above, a diffuser allows an encoding of the light field beyond just x-y coordinates and because of the random surface of a diffuser, modeling the recorded light allows us linearity in reconstruction that lends itself to well-established inverse problem definition and optimization techniques.

An important point to using diffusers and the randomness they provide is to see if this fits within the theoretical framework of compressive sensing, that has conditions on the inverse matrix. When trying to solve diffuser inverse problems with less than full rank, it is possible to recover the original vector if some conditions on a matrix are met. Such properties if satisfied such as Full Spark of matrix, and Null Space Property, both have full rank submatrices. A less computationally intensive way to calculate the sparsity that can be recovered is by computing the mutual coherence of the matrix, the combination of the measurement and representation matrix. Here is this paper we forego the formality of the mathematics and assume that our 3D objects are sparse in some domain.

A final point to consider when constructing our diffuser lensless system is to consider how a 3-D object will affect the caustic patterns on our sensor. Fig. 1 shows the effect of lateral, and depth shifts onto the sensor. Fig 2. is a top-level flow of system.

# Methods

In theory a lens less system has all the information available at the sensor. However, the problem is ill-posed. Introducing a diffuser gives structure to our point spread function, or random matrix, allowing us to solve an otherwise intractable problem. If we enforce sparsity as a prior, and non-negativity (no negative pixels), minimize for the least squares we are trying to solve the equation below

For our experiments we used a Raspberry Pi Camera V2 sensor with some modifications (see discussion section). The sensor itself is a CMOS device with 3280(H) x 2464(V) active pixel count. The platform system is a Raspberry Pi 4 Model B with HDMI ports and an interface for an external sensor. For our light source we used a flashlight with a pinhole covering the LED and black adhesive to limit output other than a pin hole to simulate our point source for calibration. Another flashlight was used to properly illuminate our object of interest.

The diffusers we used were single and double-sided ordinary scotch tape. We also constructed with two paper clips that supported a mask that fit over the CMOS sensor assembly, see Fig. 3. The entire assembly was outfitted with black opaque tape to block and limit stray light other than from our point source. All the experiments were done in as dark an environment as possible. Black tape was also used for an aperture.

A picture containing guitar

Description automatically generated

Fig. 3. Aperture, sensor, and makeshift housing to hold diffuser over sensor

We did not perform a proper calibration (see discussion for more explanation). We estimated the focal distance for our camera to just a few millimeters.

# Results

All results are using the FISTA algorithm that was outlined in the previous method section. The average computation time for reconstruction is about 30 seconds for 100 iterations. This is run on a Dell workstation that uses an Intel Core i7-8700 CPU @ 3.2Ghz with 32.0GB of RAM using Windows 10 Enterprise operating system. All code is run on PyCharm 2020.1.1 Professional Edition. Python code is downloaded and slightly modified to run on our workstation environment from. [1]

The object that was used for all experiments is a pen with the face of an animated character, see Fig 11. Fig 5. Shows our results from using double-sided tape. The algorithm for all experiments was run with at least 80 iterations using the Fast Iterative Shrinkage Thresholding Algorithm (FISTA) instead of a simple gradient. Also, for even faster speed there are options to use Alternating Direction Method of Multipliers(ADMM). Results for single sided, Antipa and double-sided response are in Figs. 5-10 respectively.

# Discussion

In the original paper [1], there is a variation where the equipment used is commodity. The construction of a system while inexpensive and convenient is not practical and stable to construct consistently with success. The author followed the instructions outlined in a tutorial and had difficulty recreating the caustic pattern point spread function.

One example of a stability and repeatable solution is the conversion of a raspberry sensor (list the type that is used) for measurements. The first problem with using a sensor like the raspberry pi sensor is that the sensor is not a flat sensor and requires that the attached lens be removed to place the diffuser a few millimeters away to produce a good caustic pattern.( see Fig. 4.) However, this requires some disassembly that is not recommended by manufacturer and if not done carefully could damage the sensor. Another difficult problem is how to determine calibration. It is challenging to hold the diffuser so close, a few millimeters away from the CMOS sensor of the raspberry pi. A final challenge is how to with a non-varying light source, calibrate the PSF at a distance from then sensor for a good reading.

As a result, with a small investment, several hundred dollars, we propose a more stable and practical solution that increases the prospect of repeated success. As a first step we construct a custom housing using a 3-d printer. Below is a diagram of the housing. (See diagram). The housing is built on top of a sensor (list part #) . The housing could stabilize the sensor and with a way to move the aperture up and down provide an easy way to perform the calibration that would yield us a good point spread function that would be manifested as an expected caustic pattern.

There are some other latent issues with the diffuser cam: Field- of-view (FoV), off-axis modeling, number of calibrations needed, resolution and finally the complexity of the object.

The first, FoV, has an angular response with incident light where the sensor will not respond. This means that when we calculate the overall Fov equation we need to account for a cutoff angle to address angles with high incidence on our diffuser. Here with “good optics”, meaning not with commodity items that we used in our experiments a FoV of +/- 42 degrees in the x -axis and +/- 30.5 degrees in y-axis was attained. For off-axis modeling, when we develop a forward model, we often assume that the diffuser is a pure phase modulation element. The light field that we approximate only then has for phase distribution only. This of course is an assumption that means that there are no effects from off-axis. A third concern with using a diffuser cam lensless system is the number of calibrations needed for good reconstruction. Here, Antipa [1] states that only one calibration is needed, yet they even recommend one at each depth of the object, which adds complexity to the overall system and makes reconstruction problematic for objects that are complex. Finally, the number of point sources of the object limit the accuracy of reconstruction. Note this is different than a regular camera where the computation is not a function of the object. Antipa [1] and Cai [8] both show the effect of increasing the complexity of and object and increase of the condition number for an increase in angular sampling.

# Conclusion

Following the guide outlined in Antipa [1], we tried to build with the same components that were outlined in their tutorial. We listed two of the challenges encountered, one was modifying the sensor and the second was performing calibration. We were not able to get a satisfactory caustic pattern that would allow us

with some accuracy to recover the object. Also, with both of our chosen diffusers, single sided tape, and double-sided tape, we had the same point spread function on our sensor.

There are three things that we would like to expand on after this project. The first is <We are currently setting up a lab to replicate the work of Antipa…..>< Here are some specs from the system that we are trying to build >.

The second item that we hope to explore is to model PSFs with a formula like Jin[X} and also look at off -axis performance. The paper from( reference <JIN> outlines a method that can help with avoiding the use of extensive calibration in order to provide better reconstruction. The idea is based on using Fourier optics first to express the system PSF as a function with only one unknown variable, which is the diffuser phase distribution. Then JIn et al. proceeds to estimate the diffuser phase by using laser beam shaping theory. Using an architecture with a Fourier lens between the diffuser and the sensor, and that fact that a far-field distribution of laser beams is proportional to the Fourier Transform of the near-field distribution, the phase of the diffuser corresponds to the phase of the diffuser which then corresponds to the near-field component, retrievable by phase retrieval methods( here we make reference to JIn’s paper and use ref 20-22}

The third area that we would like to expand on is how the models are built around light field fundamentals. For example, our diffuser cam using the ideas outlined by Antipa[1] is built by using ray optics. Here we note the main limitation of this approach is that it assumes that the diffuser to sensor distances is small. For many applications this is not a good assumption. An alternative way would use wave-optics that better model diffractive effects. Here the Wigner distribution, originally developed to study quantum mechanics, can be used to model the light field without the diffuser to sensor constraint( see reference for wigner)

The overall paradigm of using a diffuser to capture 4-D plenoptic information in a 2-D sensor in engineering terms means less samples and an overall reduction in the complexity of the acquisition system. This translates into 1) less power, 2) less hardware ,3) simpler algorithms for acquiring (directly proportional to hardware), 4) less speed for sampling (even non-uniform sampling), 5) less bandwidth, and 6) less storage and a general ability of the engineer to trade-off performance given the “compactness” of sensing.

Further developing diffuser light field modeling to the point where complex objects can be reconstructed with robust and repeatable results will allow the realization of many new and interesting applications.

##### References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the

first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*
2. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
3. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
4. K. Elissa, “Title of paper if known,” unpublished.
5. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
6. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
7. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**

Text, letter

Description automatically generated